

2017 FLASH FLOOD AND INTENSE RAINFALL EXPERIMENT – FV3-GFS PERFORMANCE AND FEEDBACK

Benjamin Albright – WPC-HMT

Joshua Kastman – WPC-HMT

Mike Bodner – WPC-DTB/HMT

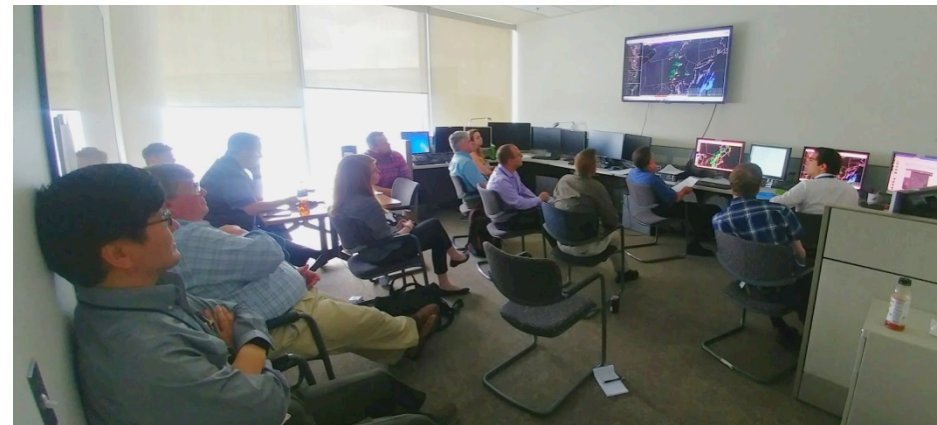
Weekly FV3-GFS Technical Meeting – August 21, 2017

Outline

- Overview of the 2017 Flash Flood and Intense Rainfall (FFaIR) Experiment.
 - Goals
 - Forecast Activities
 - Other Experimental Guidance
- Subjective Verification
- FV3-GFS Examples
- MODE Analysis
- Objective Statistics
- Day 8-10 Experiment

2017 FFaIR Experiment Goals

- 5th Annual FFaIR Experiment.
- Held over four weeks from June 19 – July 21, 2017.
- Major goals included:
 - Evaluate the utility of high resolution convection-allowing deterministic models and other guidance for flood forecasts at longer time ranges (Day 2-3).
 - Explore changes to WPC's operational Excessive Rainfall Outlook.

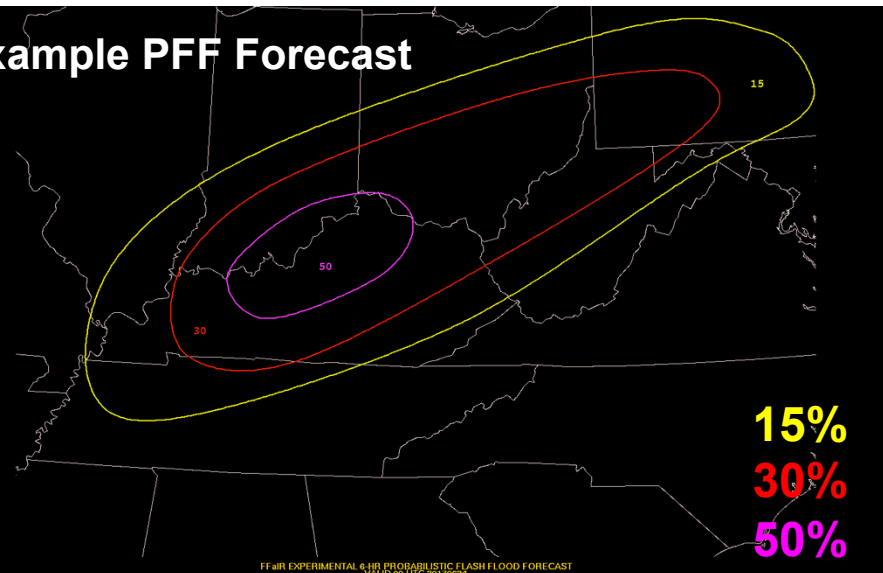


FFaIR Experiment participants review a morning forecast with real time data in the late afternoon.

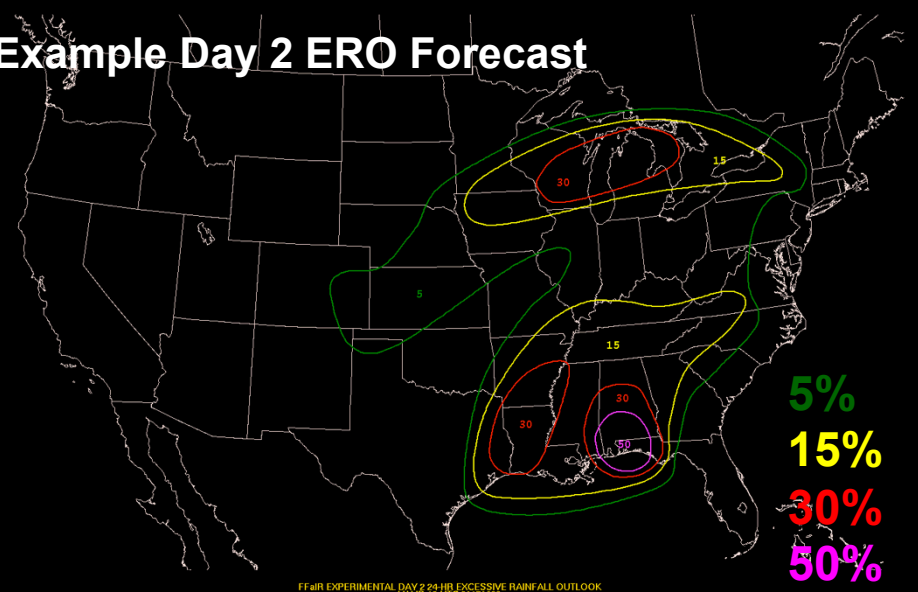
2017 FFaIR Forecast Activities

- Six hour Probability of Flash Flooding Forecast (PFF)
 - Issued by 1545Z, valid 18-00Z same day.
- Day 2 Excessive Rainfall Outlook (ERO)
 - Probability of “flooding rains” occurring, valid 12Z-12Z, 36-60 hour forecast.
- Day 2 Flood Watch
- Day 3 ERO
 - Probability of “flooding rains” occurring, valid 12Z-12Z, 60-84 hour forecast.

Example PFF Forecast



Example Day 2 ERO Forecast

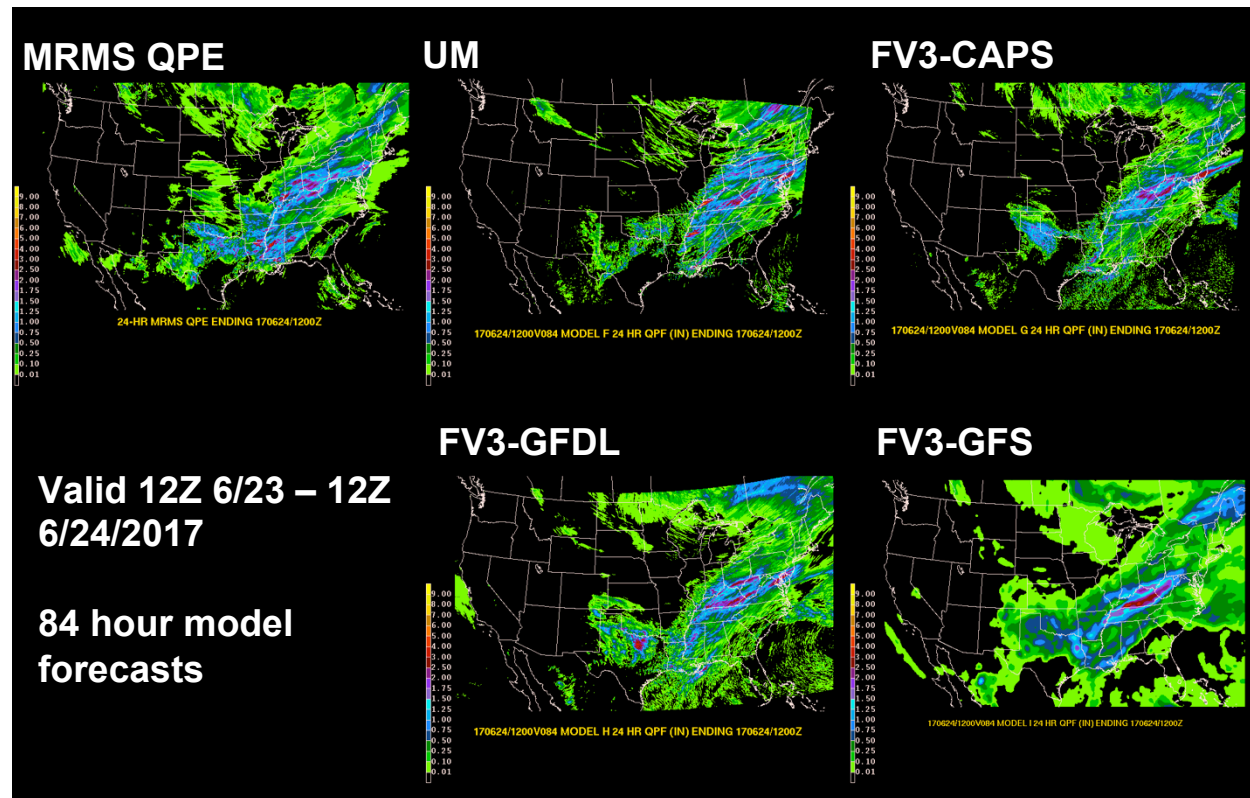


Select Day 2/Day 3 2017 FFaIR Model Guidance

Provider	Model	Resolution	Forecast Hours	Notes
Met Office	UM	2.5 km	84 hours <i>(for experiment)</i>	Convection allowing; sub-CONUS domain, 70 vertical levels up to 40km; UKMet boundary conditions/initialization
OU/CAPS	FV3-CAPS	3 km	84 hours <i>(for experiment)</i>	Using the FV3 dynamical core and Thompson microphysics
GFDL	FV3-GFDL	3 km	84 hours <i>(for experiment)</i>	Using the FV3 dynamical core and GFDL microphysics
EMC	FV3-GFS	13 km	84 hours <i>(for experiment)</i>	3D hydrostatic dynamical core; vertically Lagrangian; GFS analyses initialization/physics

Verification

- During subjective verification, model QPF was verified against radar-based MRMS QPE.
- Participants assigned subjective scores of 1 (very poor) to 10 (very good) for each model.
- FV3-GFS was evaluated at the Day 3 (84 hour) time period and will be the focus of the following results/slides.

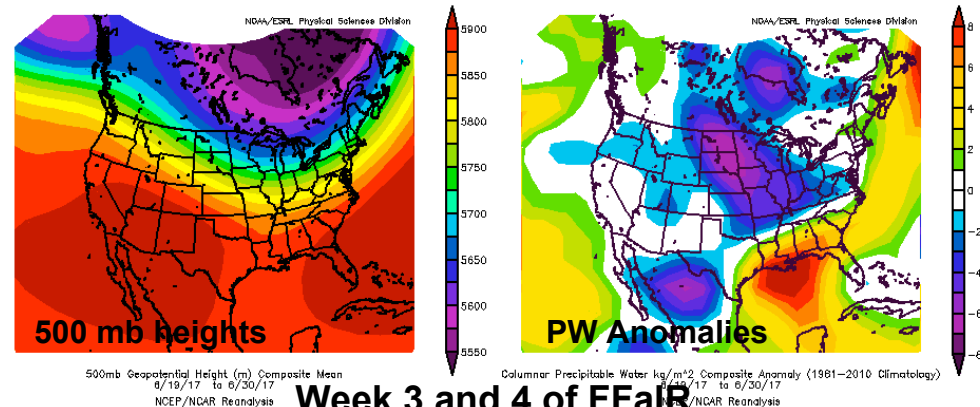


An example of how Day 3 model QPF was presented to participants for subjective evaluation.

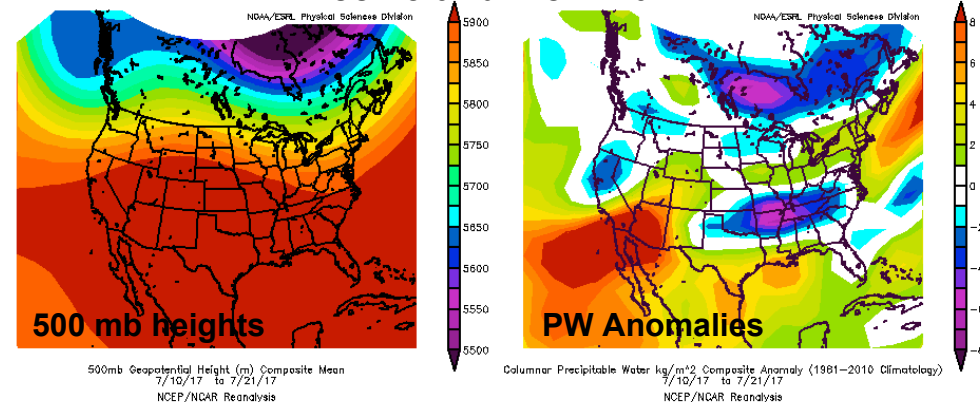
2017 FFaIR Synoptic Overview

- Week One – Dominated by Tropical Storm Cindy and frontal boundary.
- Week Two – Scattered diurnal thunderstorms/MCS activity late week.
- Week Three/Four – Dominated by Upper Midwest MCS activity and Southwest monsoon moisture.

Week 1 and 2 of FFaIR

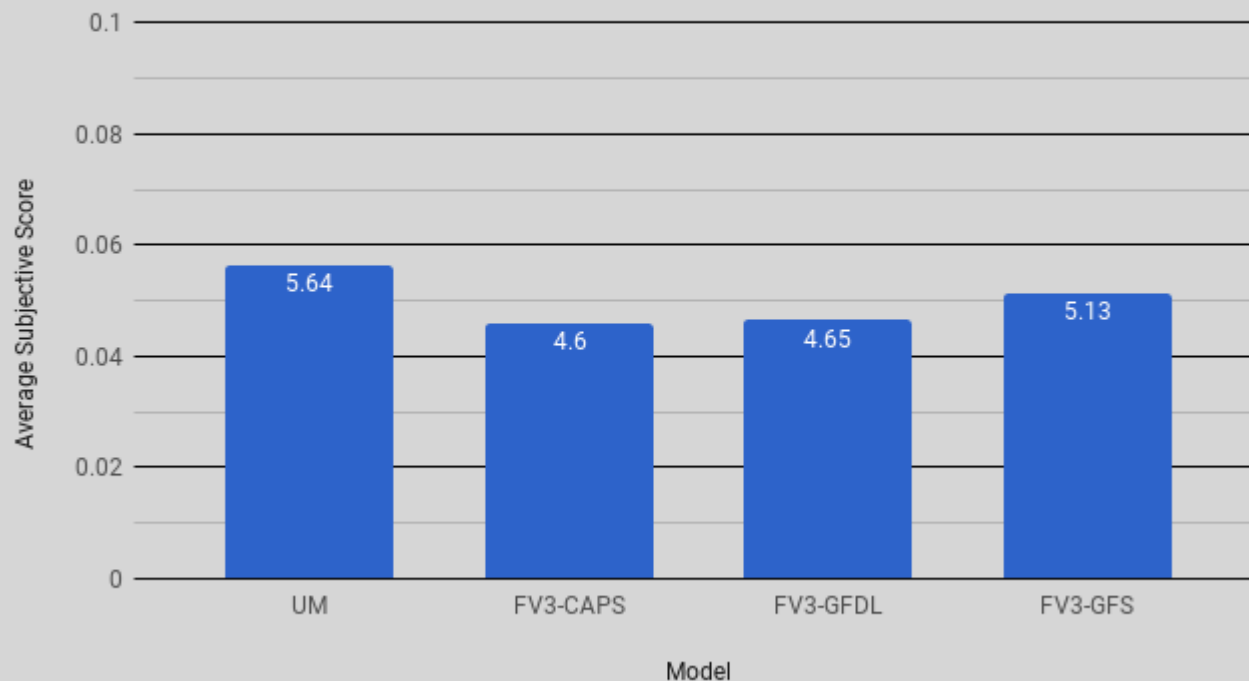


Week 3 and 4 of FFaIR



FV3-GFS Subjective Verification

Average Subjective Scores During 2017 FFaIR for Day 3 Deterministic 24 HR QPF



Total Scores Over Four Weeks:

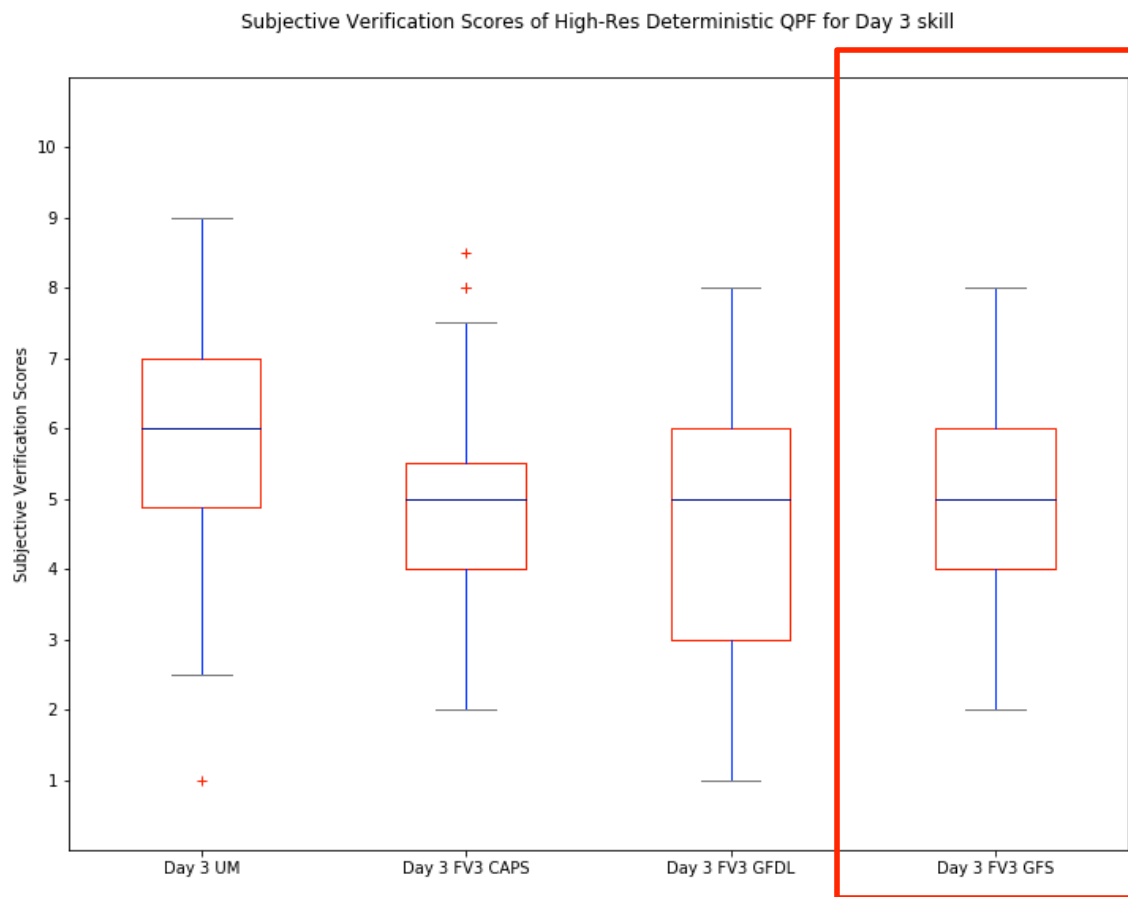
UM-----> 120

FV3-CAPS--> 127

FV3-GFDL--> 129

FV3-GFS----> 151

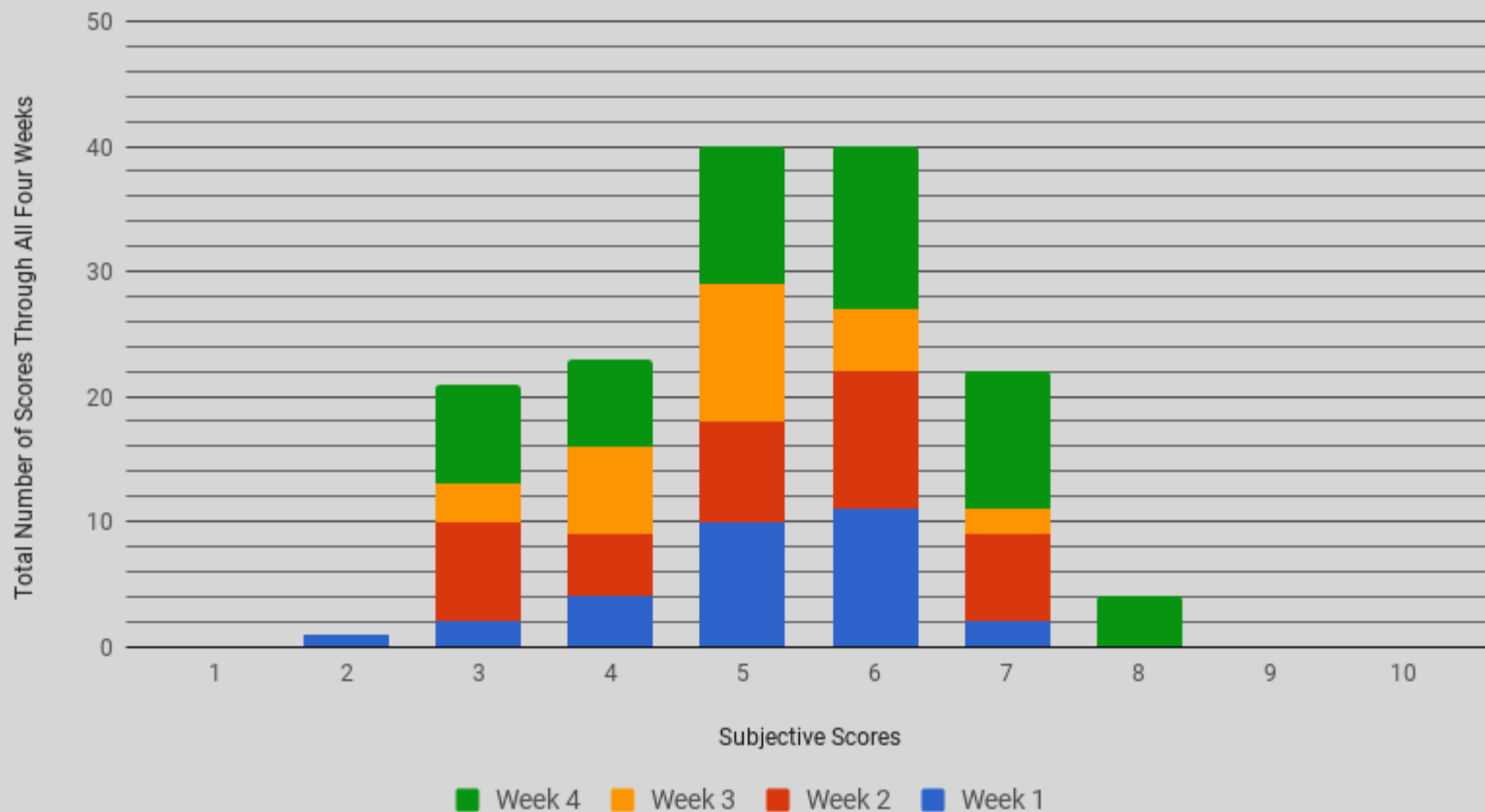
FV3-GFS Subjective Verification



Box plot of all the subjective scores for the Day 3, 84 hour forecasts for 24 hour QPF from the UM, FV3-CAPS, FV3-GFDL, and FV3-GFS over the whole experiment. Red plus symbols indicate outliers.

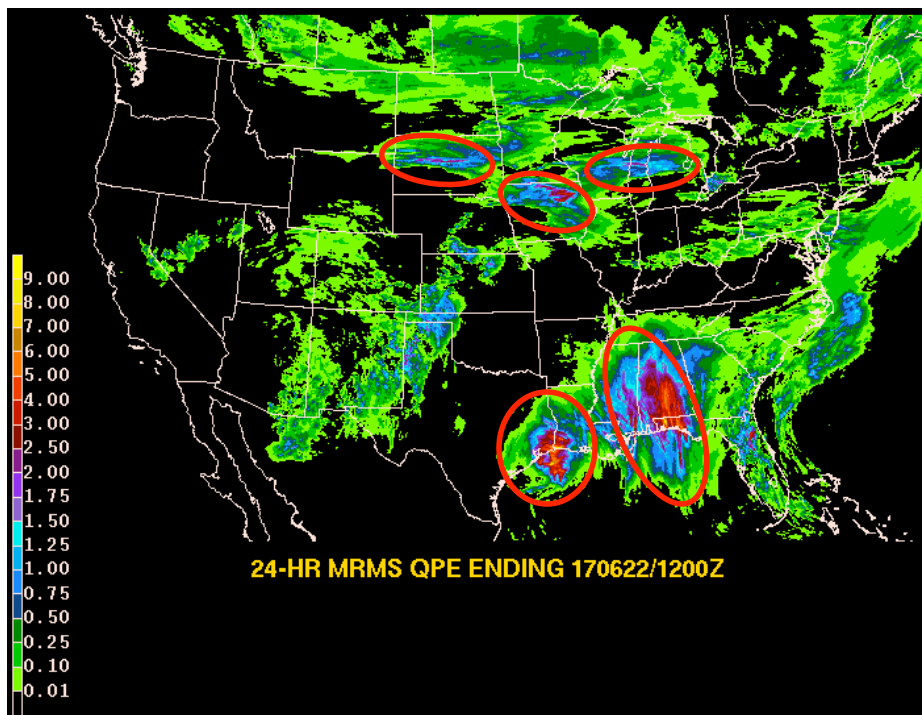
FV3-GFS Subjective Verification

FV3-GFS Distribution of Individual Subjective Scores Through FFaIR 2017

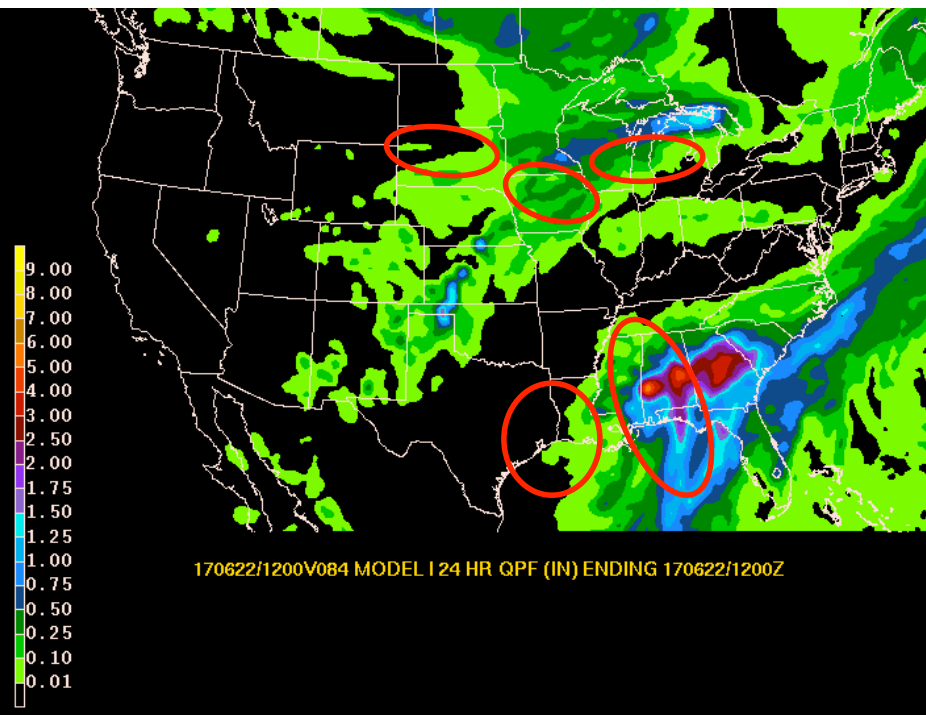


FV3-GFS FFaIR 2017 Examples

24 HR MRMS QPE



24 HR FV3-GFS QPF (84 HR FCST)



VALID 12Z 6/21 – 12Z 6/22/2017

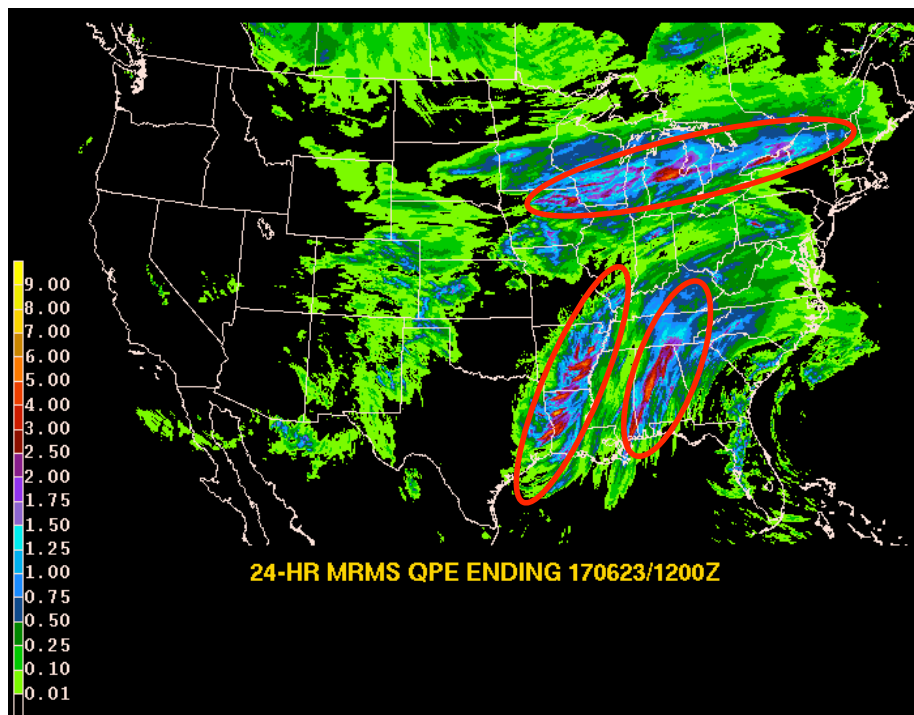
Average subjective score → 5/10

Comments:

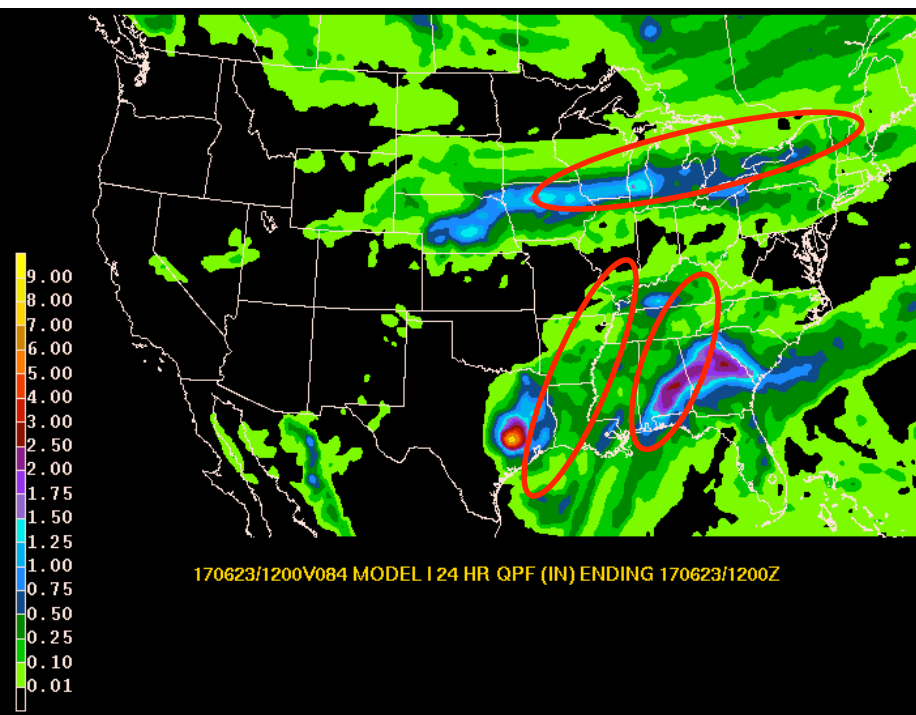
- *Did well with Midwest shortwave, some displacement issues, lower magnitude.*
- *Brings TC onshore too fast, too far east.*

FV3-GFS FFaIR 2017 Examples

24 HR MRMS QPE



24 HR FV3-GFS QPF (84 HR FCST)



VALID 12Z 6/22 – 12Z 6/23/2017

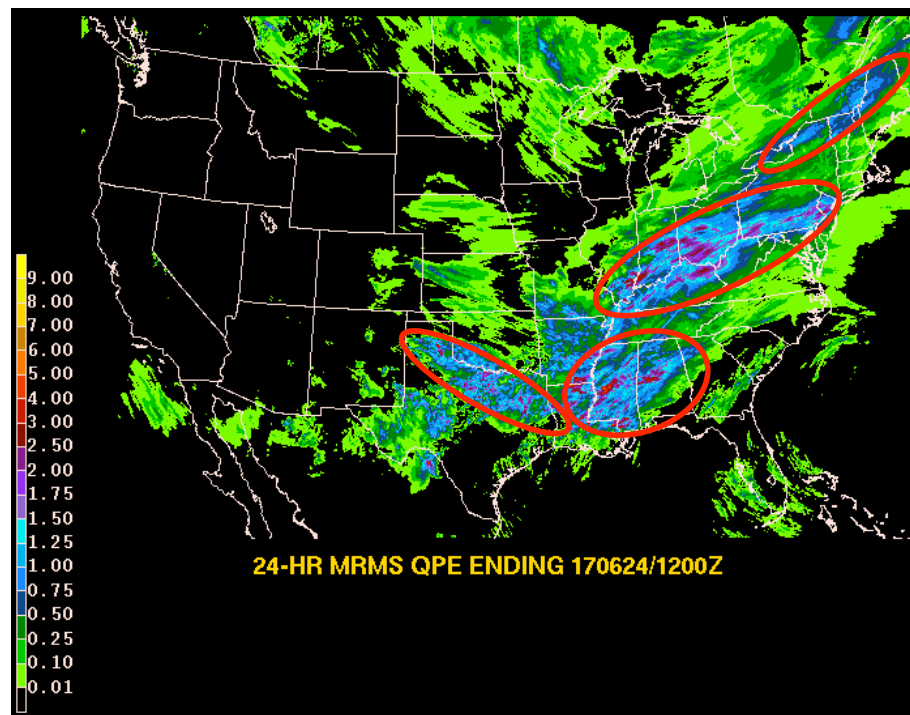
Average subjective score → 4/10

Comments:

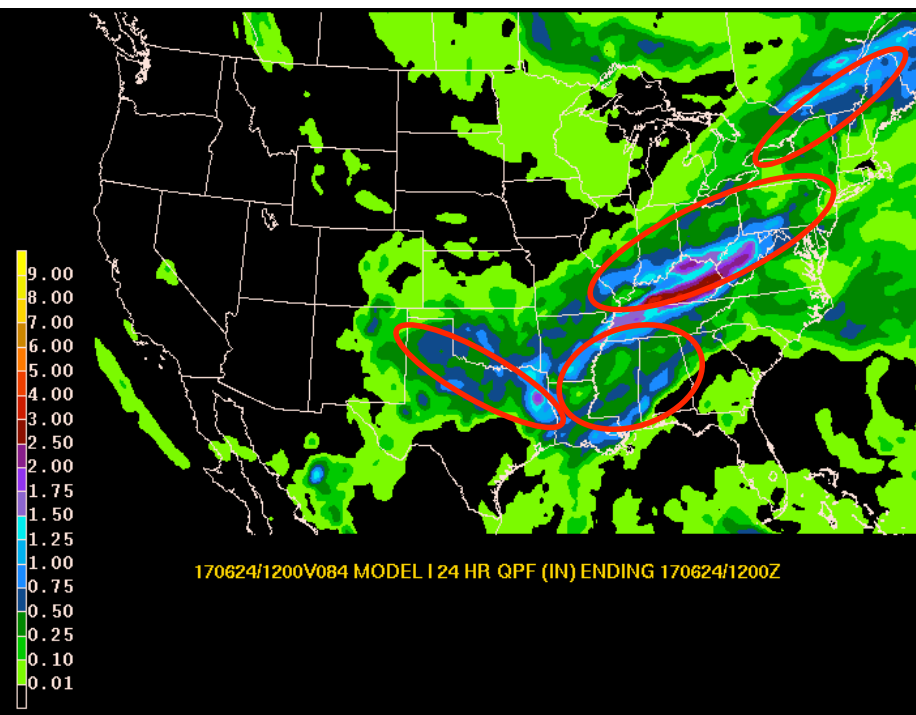
- Midwest frontal boundary location good, magnitude too low.
- Eastern location of TC wrong, too high magnitude.
- Stripe of rain in Alabama/Georgia better.

FV3-GFS FFaIR 2017 Examples

24 HR MRMS QPE



24 HR FV3-GFS QPF (84 HR FCST)



VALID 12Z 6/23 – 12Z 6/24/2017

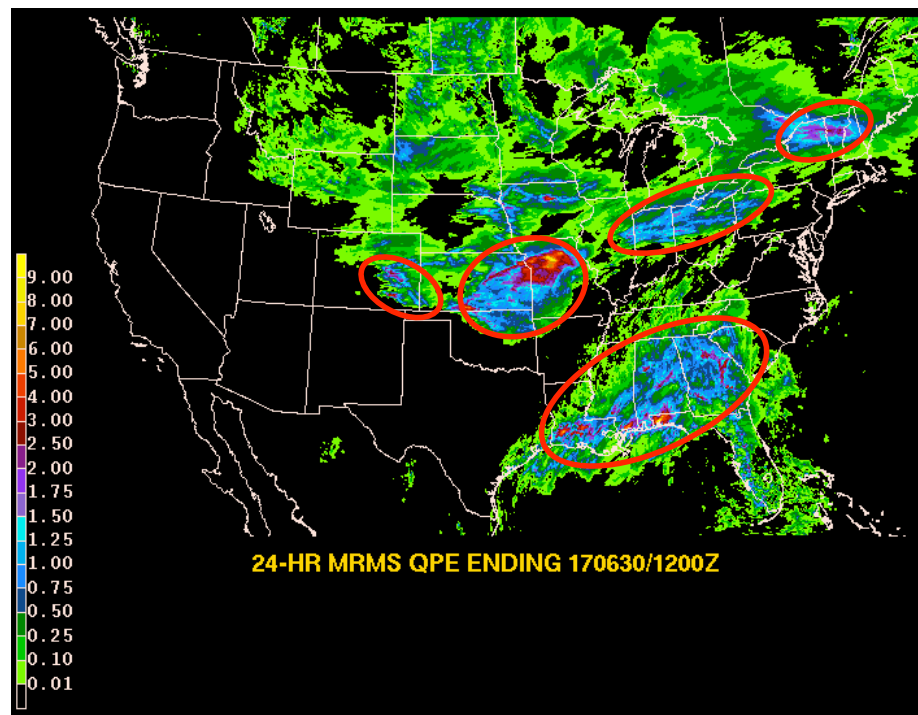
Average subjective score → 6/10

Comments:

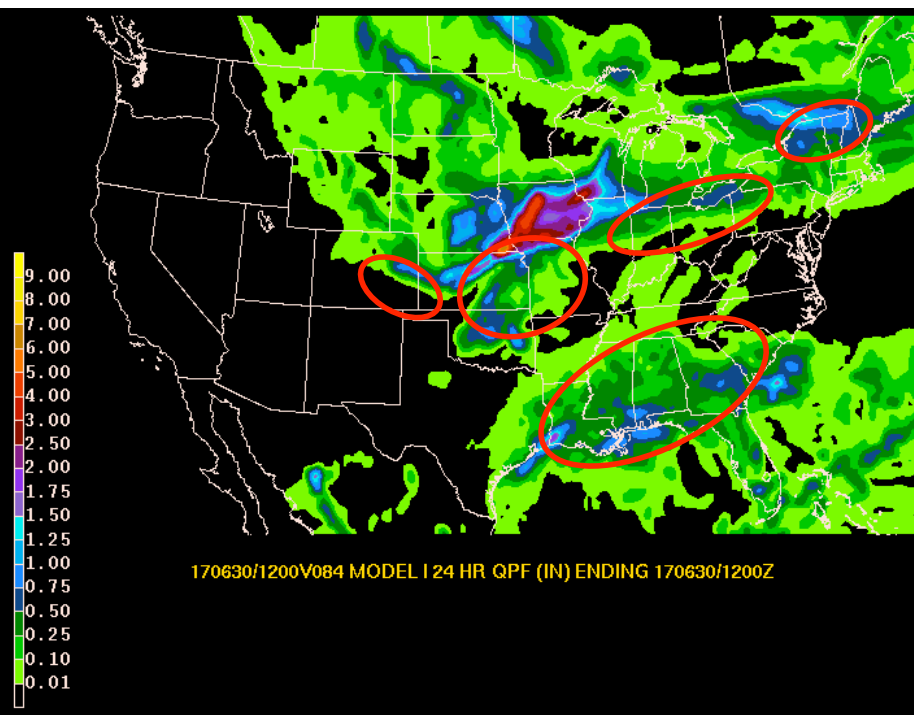
- Has general idea of the areas of heavier precipitation. Stripe through Ohio River Valley a little too far south.
- Other areas a little low in magnitude, especially over AL/MS.

FV3-GFS FFaIR 2017 Examples

24 HR MRMS QPE



24 HR FV3-GFS QPF (84 HR FCST)



VALID 12Z 6/29 – 12Z 6/30/2017

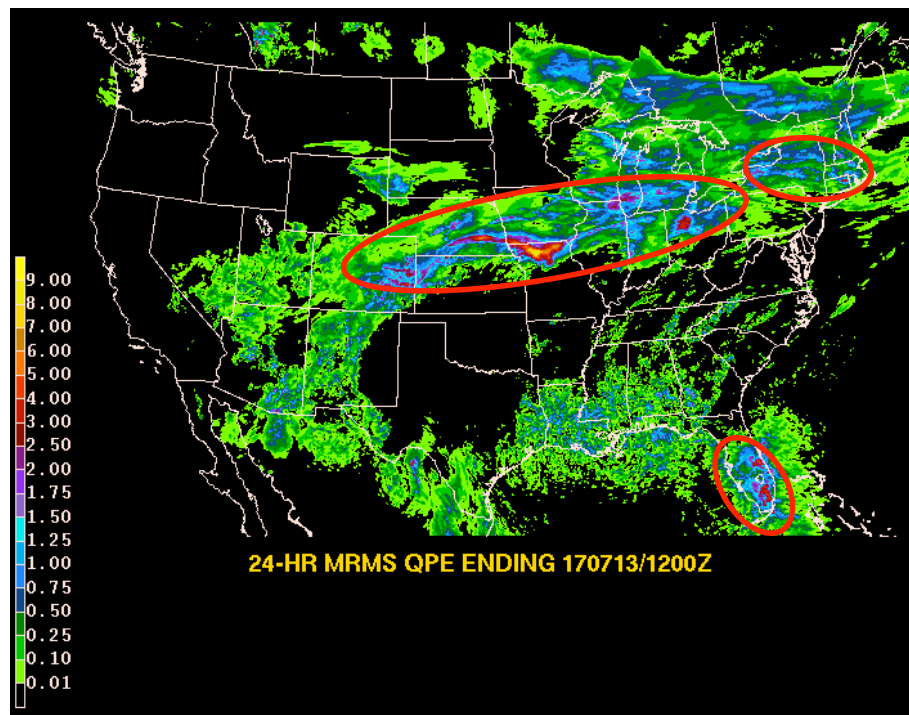
Average subjective score → 5/10

Comments:

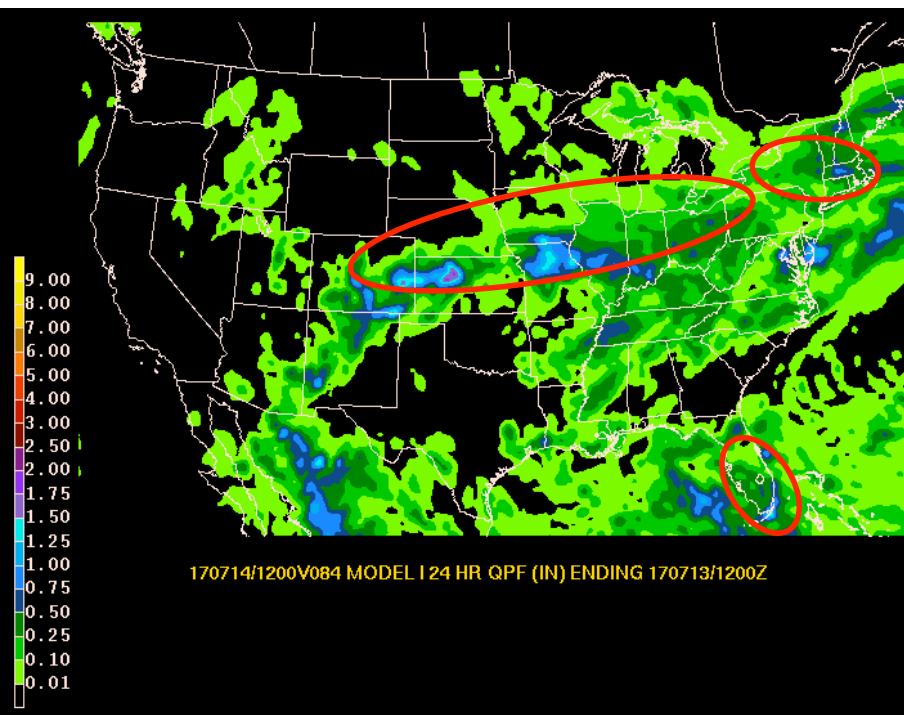
- *Maximum in Plains displaced to the north.*
- *New England area pretty good.*
- *Hints at stripe in Colorado.*
- *Southeast magnitude is low.*

FV3-GFS FFaIR 2017 Examples

24 HR MRMS QPE



24 HR FV3-GFS QPF (84 HR FCST)



VALID 12Z 7/12 – 12Z 7/13/2017

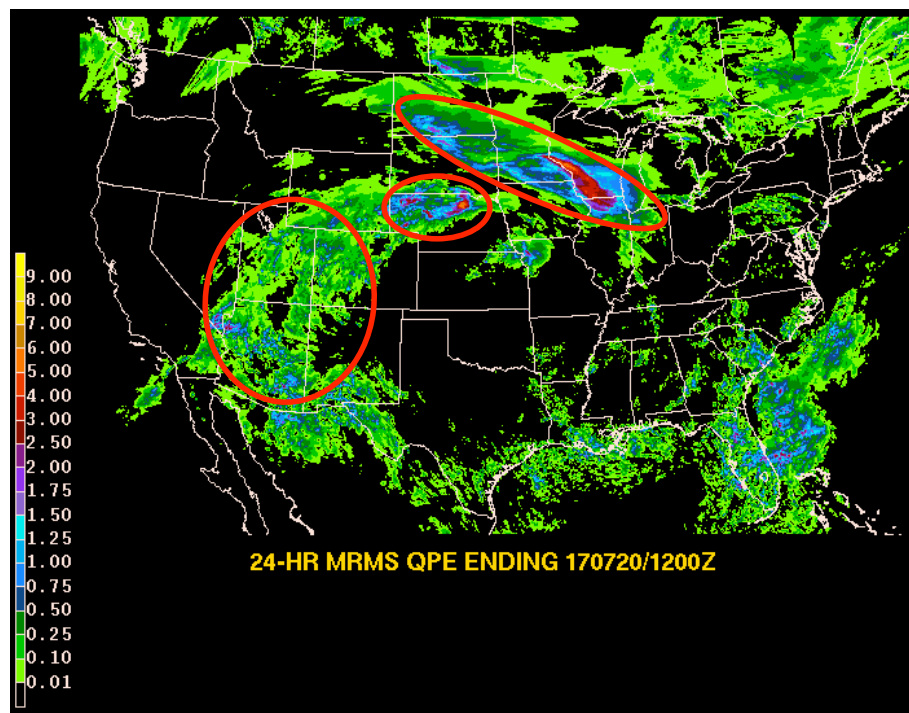
Average subjective score → 5/10

Comments:

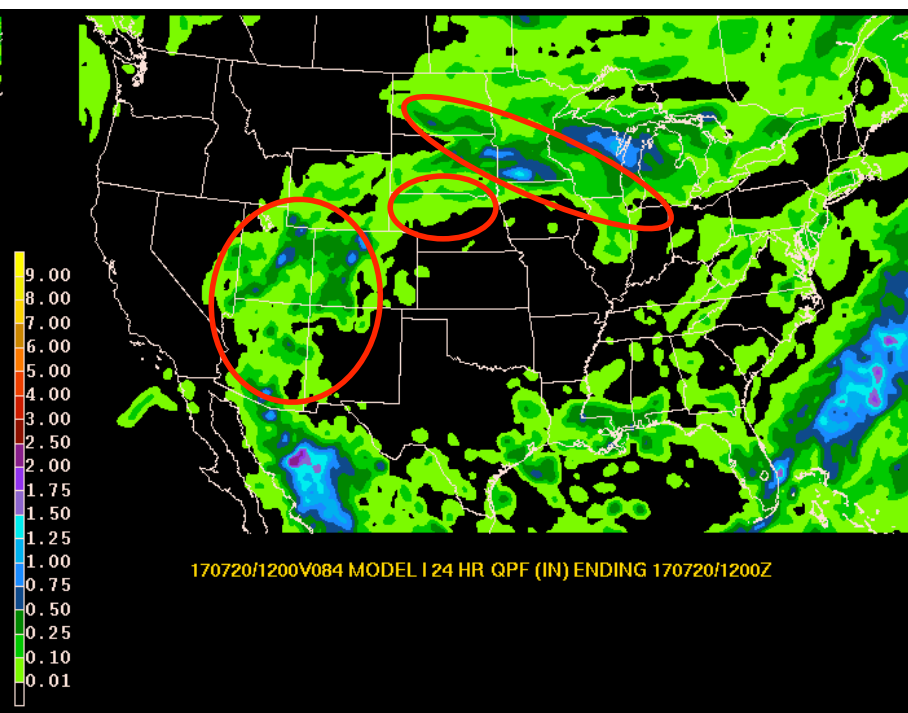
- Good placement in Missouri and western Kansas/ eastern Colorado. Amounts low.
- Broad spatial pattern makes it hard to focus on areas. Much of Southeast has too much precip.

FV3-GFS FFaIR 2017 Examples

24 HR MRMS QPE



24 HR FV3-GFS QPF (84 HR FCST)



VALID 12Z 7/19 – 12Z 7/20/2017

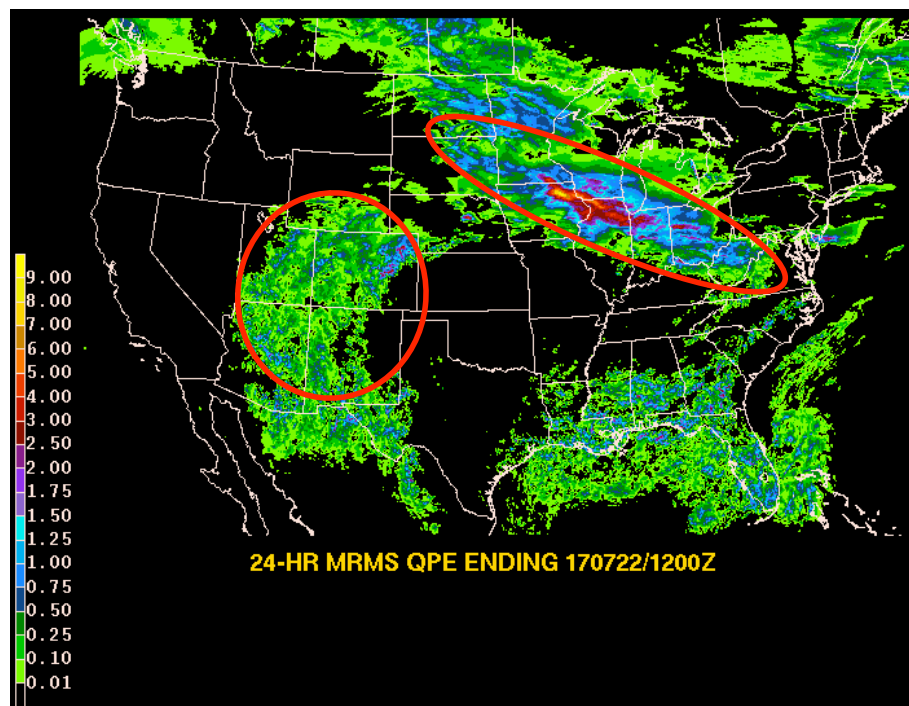
Average subjective score → 5/10

Comments:

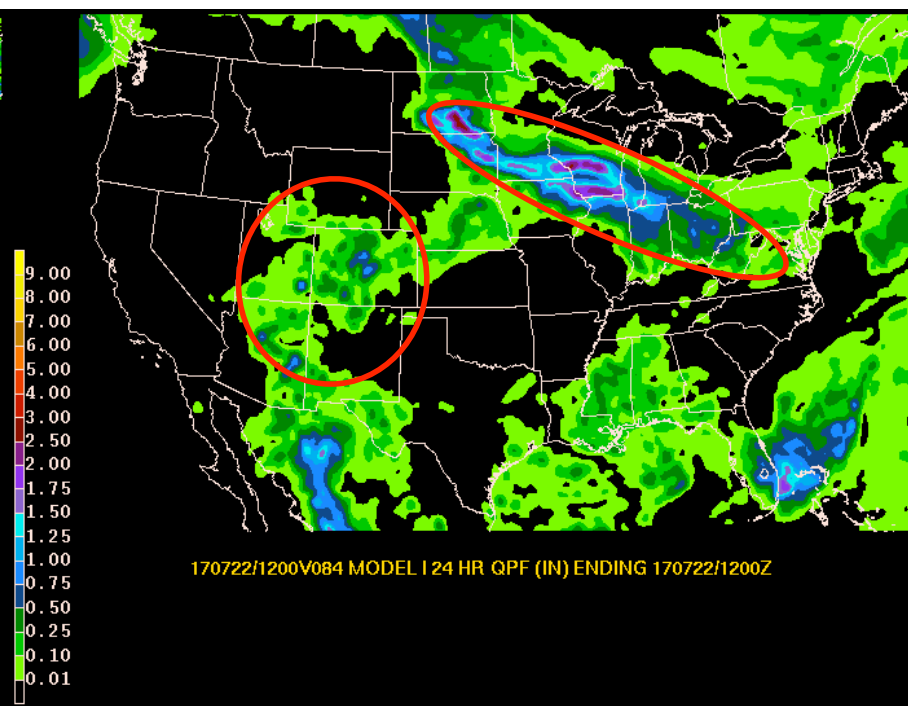
- Wisconsin maximum too far north, underdone.
- Southwest coverage spatially good.

FV3-GFS FFaIR 2017 Examples

24 HR MRMS QPE



24 HR FV3-GFS QPF (84 HR FCST)



VALID 12Z 7/21 – 12Z 7/22/2017

Average subjective score → 7/10

Comments:

- Main axis well represented in Central Plains. A little too far north, too heavy in North Dakota.
- Southwest and Southeast regions both have good spatial coverage.

Conclusions

- Subjectively, the FV3-GFS had an average score *higher* than the higher resolution FV3-CAPS and FV3-GFDL for the Day 3 period and a half point lower than the Unified Model from the Met Office.
- Comments were generally favorable regarding the spatial distribution of precipitation.
- Precipitation maximums were often too low, although forecasters did adjust to this somewhat.
- Most common displacement seen subjectively: precipitation maximums that were too far north, especially for MCS activities in the Great Plains.

FFaIR Objective Verification

FV3-GFS Objective Verification

- 60 HR & 84 HR QPF verified against Stage IV QPE
 - 00Z forecast cycle used
 - Both QPF and QPE regridded to a common 5km grid
 - CONUS mask applied to common grid
 - Thresholds of 0.5", 1.0", 2.0", 4.0" and 6.0" investigated
- MODE
 - Grid stats harvested from MODE CTS
 - Circular convolution radius of 3 grid squares used
 - Double thresholding technique applied
- MODE Analysis
 - Summary of all forecasted vs. observed shapes throughout experiment
 - Describes centroid distance, angle, and interest

FV3-GFS Objective Verification Observations

- Several instances of Day 3 outperforming Day 2 were observed
 - 6 times in 23 comparable forecasts at the 0.50" threshold
 - 9 times in 23 comparable forecasts at the 1.00" threshold
- Struggled with scattered convection
- Excelled in MCS & continuous precipitation situations
 - Many of the high resolution simulations did not yield a continuous precipitation shield
 - "Pointillism" appearance often observed
- Major skill drop off between 0.50" and 1.00"

FV3-GFS Performance Diagrams

Roebber Performance
Diagram

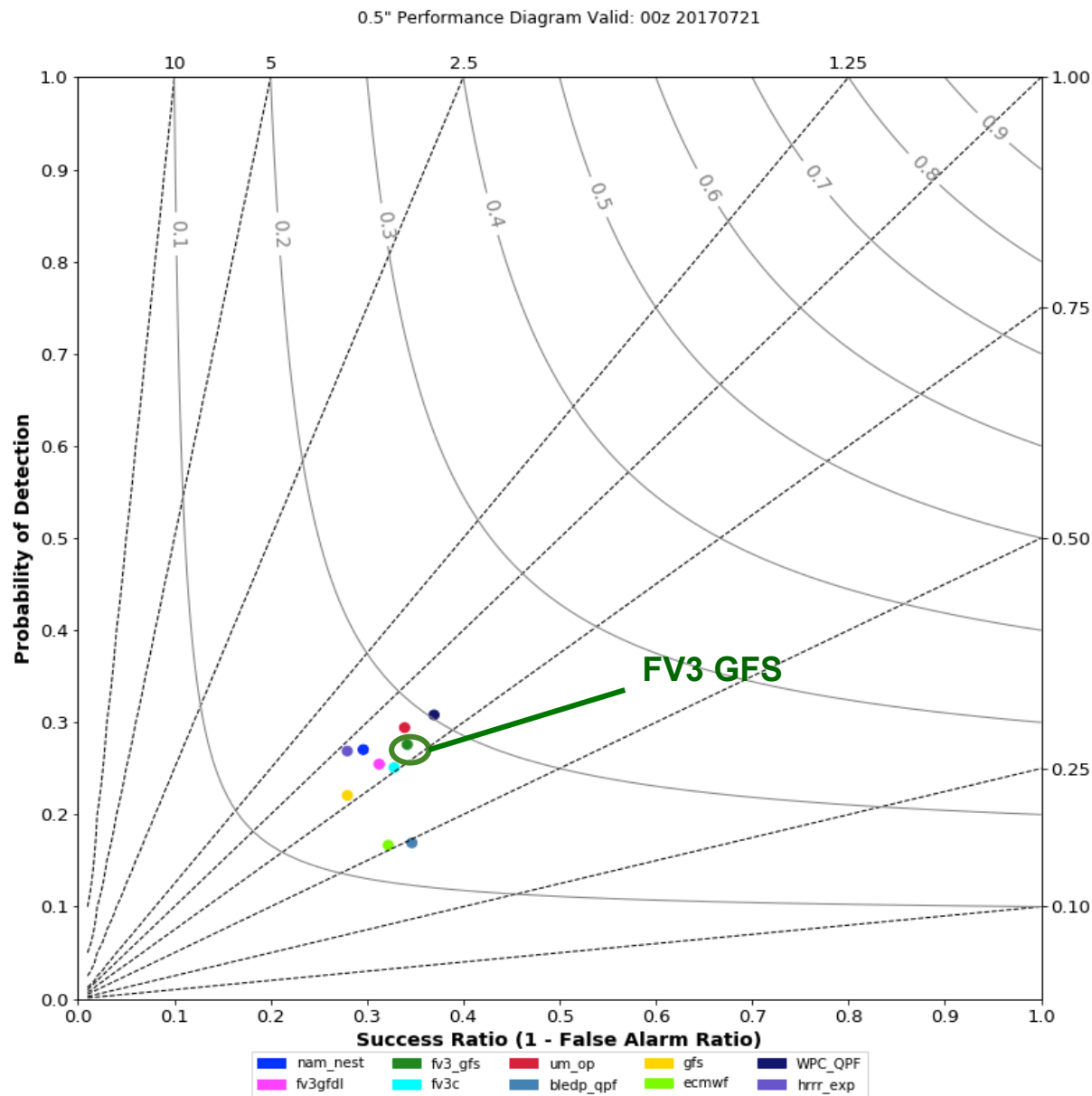
Running Mean
Throughout Experiment
(June 17 - July 21)

60 HR Forecast

0.50" Threshold

FV3-GFS CSI: 0.1845

3rd Best behind UM and
WPC (10 total)



FV3-GFS Performance Diagrams

Roebber Performance
Diagram

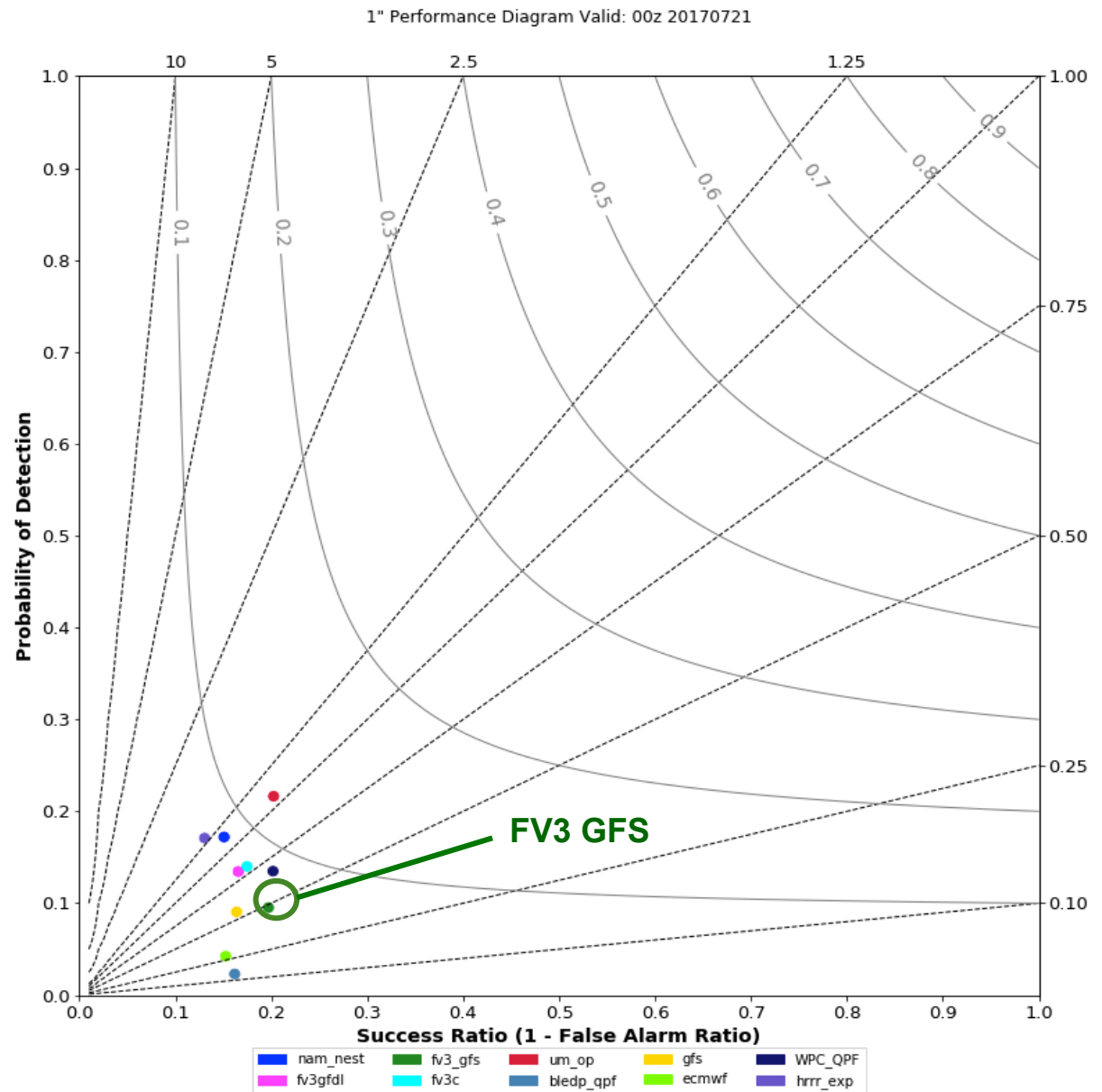
Running Mean
Throughout Experiment
(June 17 - July 21)

60 HR Forecast

1.00" Threshold

FV3-GFS CSI: 0.0635

7th out of 10 models



FV3-GFS Performance Diagrams

Roebber Performance
Diagram

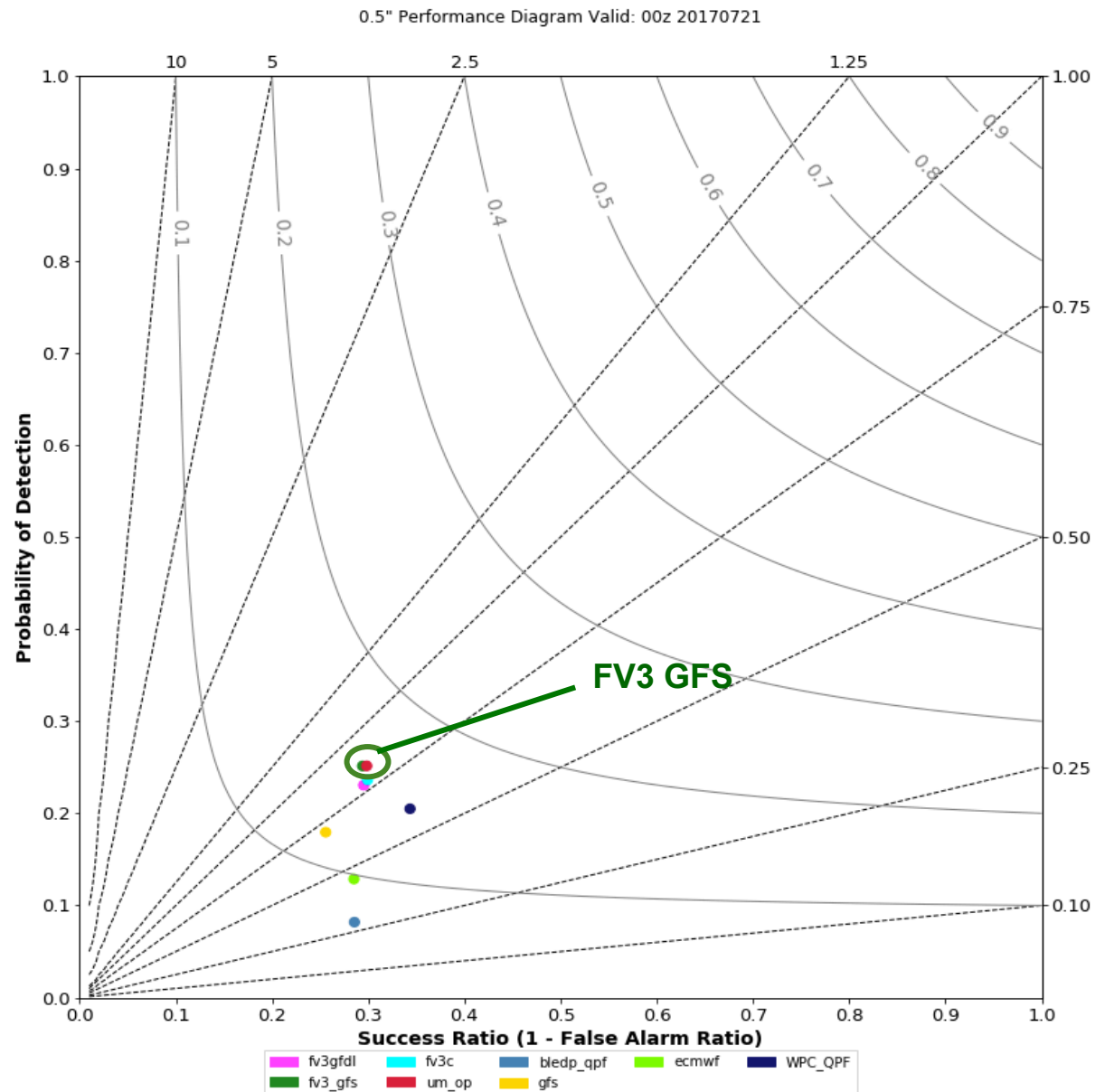
Running Mean
Throughout Experiment
(June 17 - July 21)

84 HR Forecast

0.50" Threshold

FV3-GFS CSI: 0.1588

2nd Best behind UM
(8 total models)



FV3-GFS Performance Diagrams

Roebber Performance
Diagram

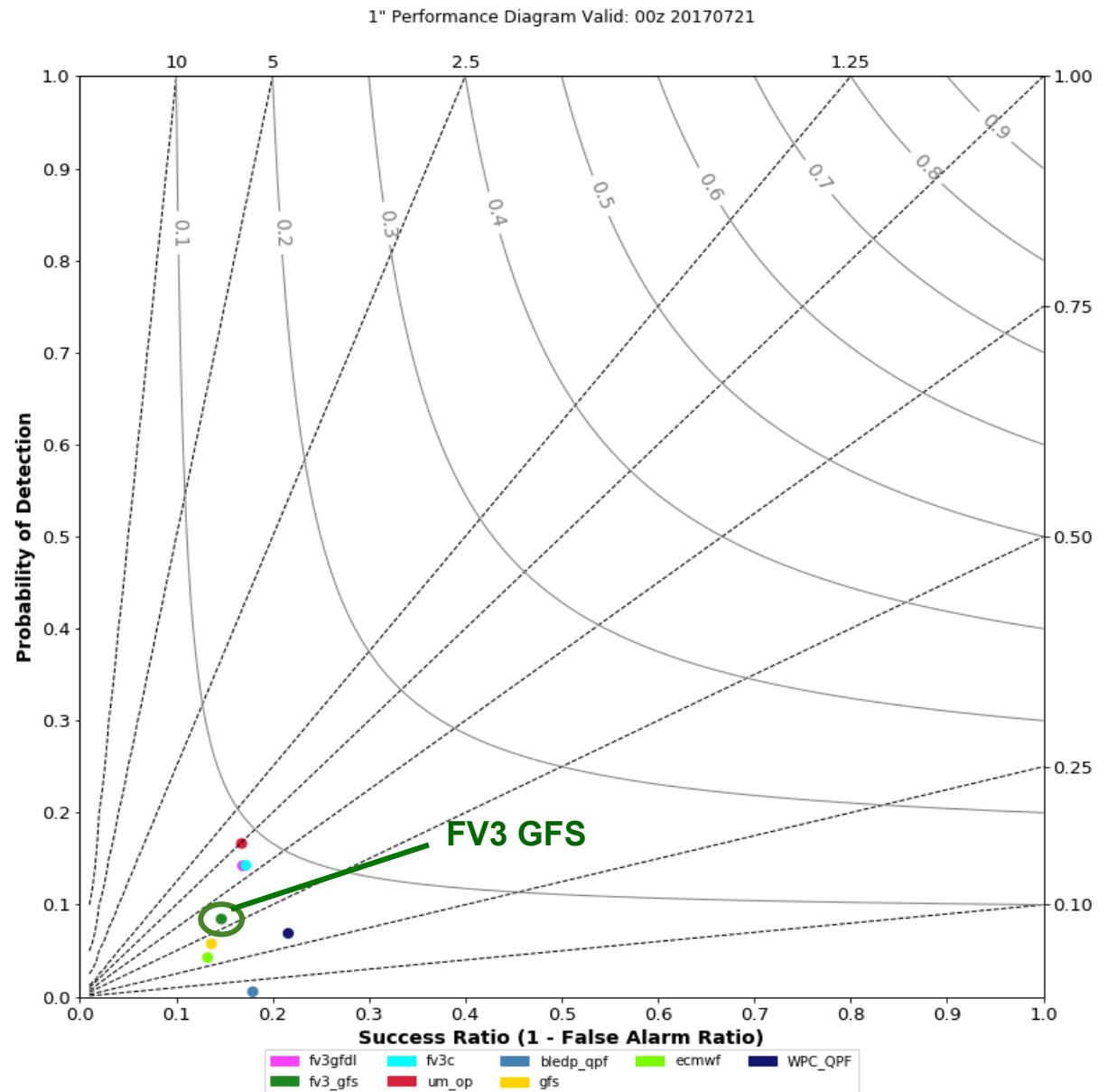
Running Mean
Throughout Experiment
(June 17 - July 21)

84 HR Forecast

1.00" Threshold

FV3-GFS CSI: 0.0508

4th out of 8 models



FV3-GFS MODE Examples

FV3_GFS 0.5° Day 2 24 HR QPF | Initialized: 2017062900 | Valid: 2017063012Z - 20170701 12Z

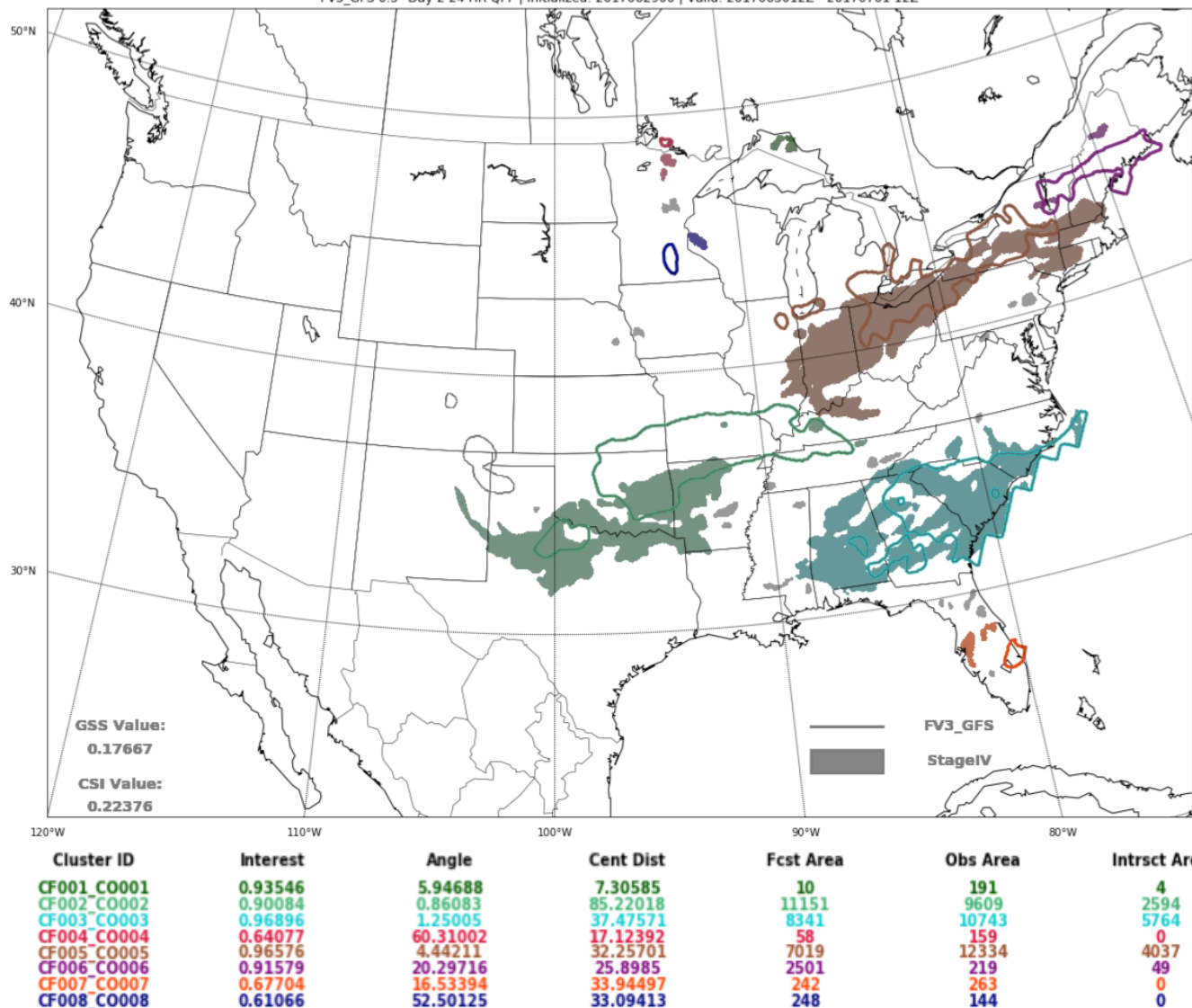
60 HR Forecast valid
12 UTC 01 July 2017.

0.5" Threshold

Each color represents
a matched forecast
object and observed
object.

Gray lines indicate
unmatched forecast

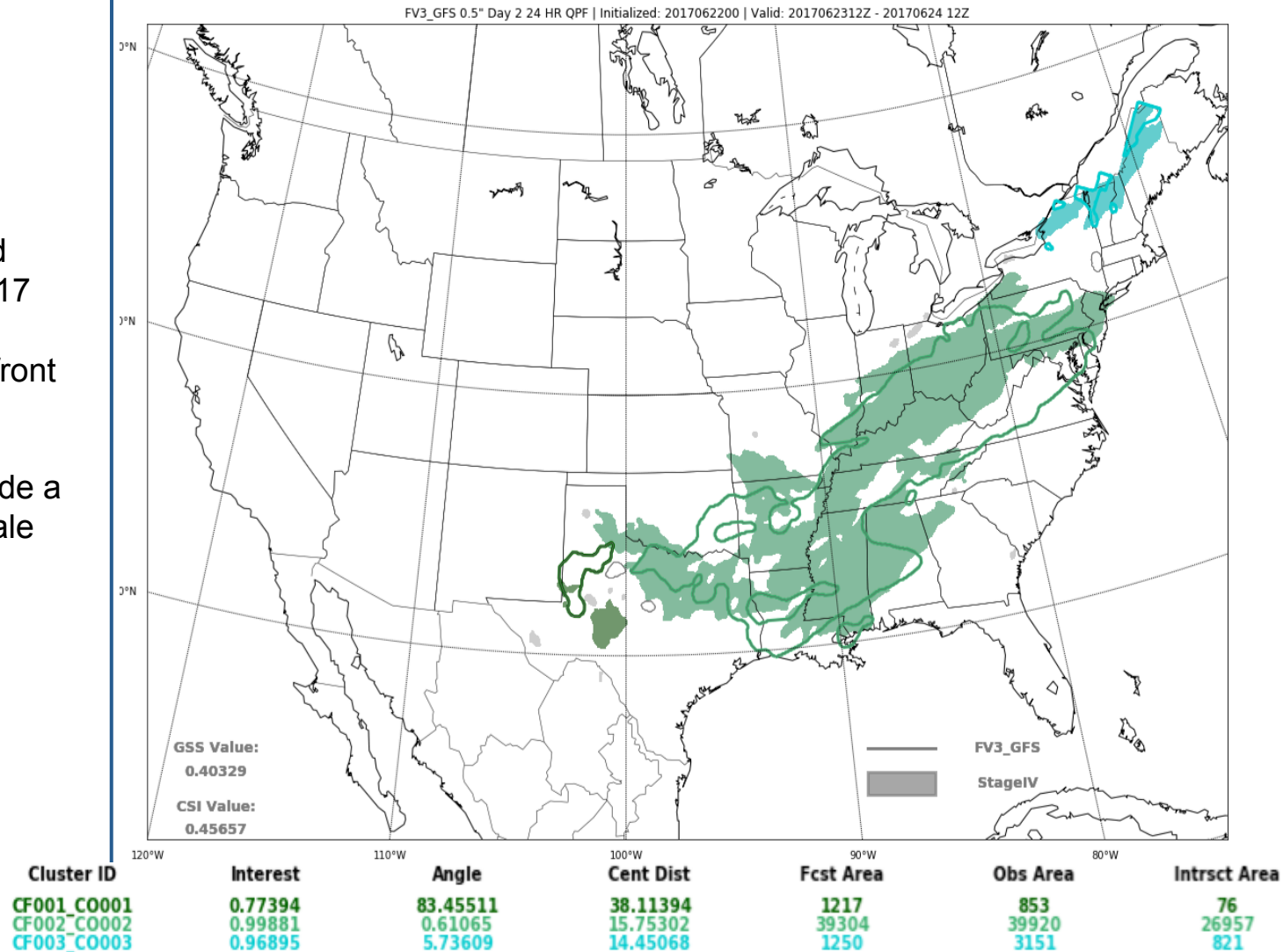
Gray colorfill indicates
unmatched
observations



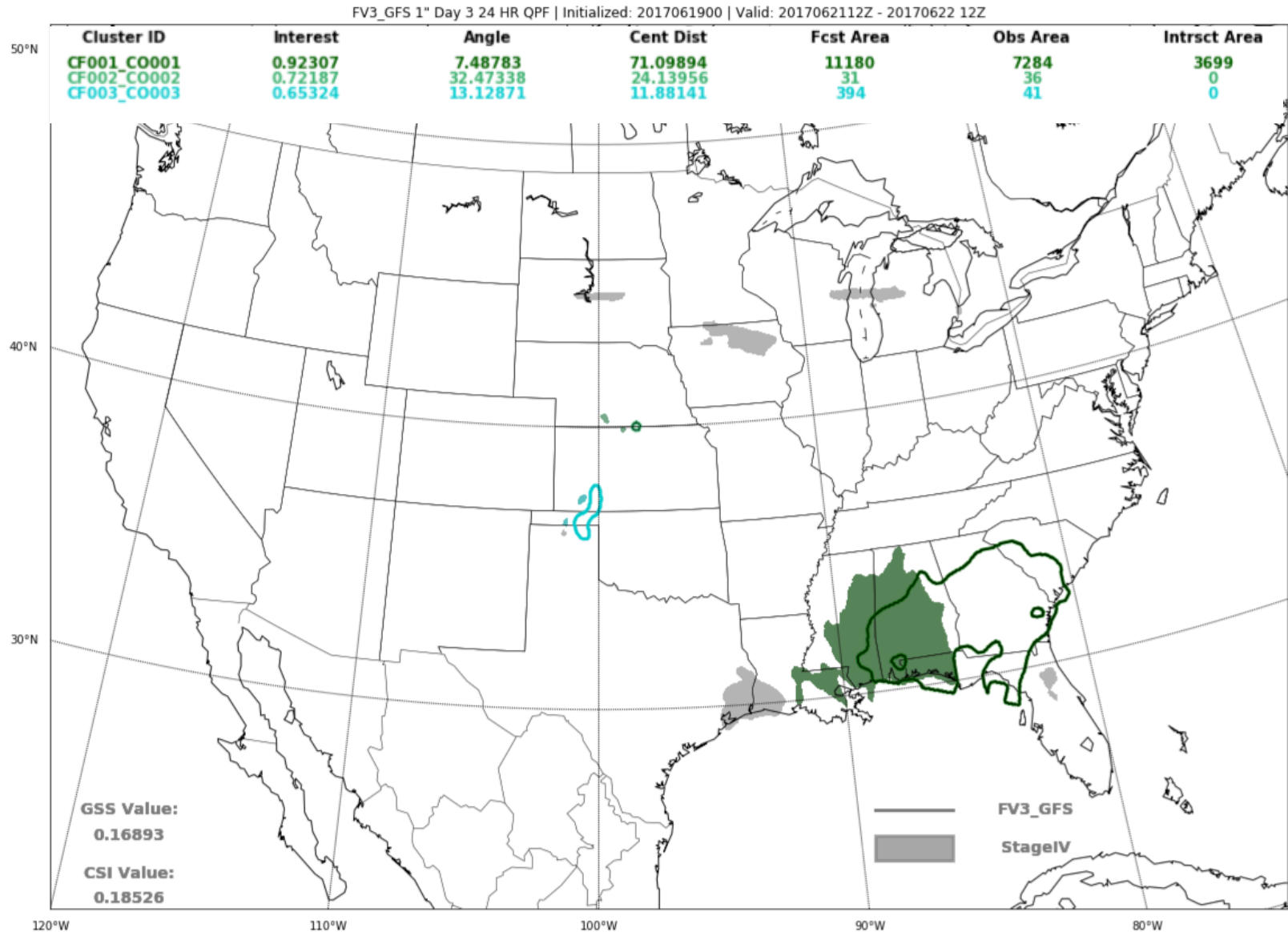
FV3-GFS MODE Examples

Best Performance:
60 HR Forecast valid
12 UTC June 24, 2017

Synoptic scale cold front
combines with the
remnants of Tropical
Storm Cindy to provide a
well forced, large scale
heavy rain area.

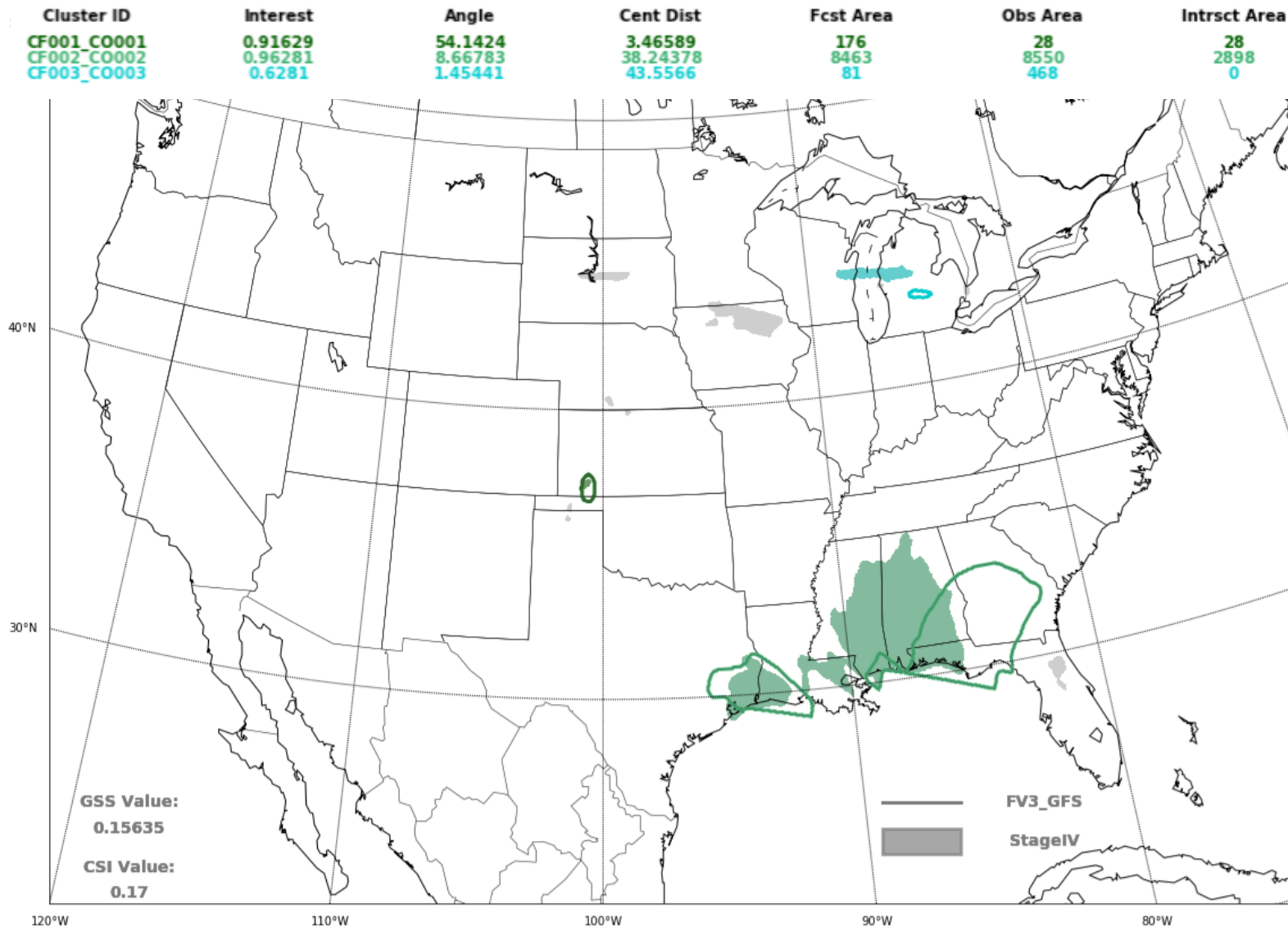


FV3-GFS MODE Examples

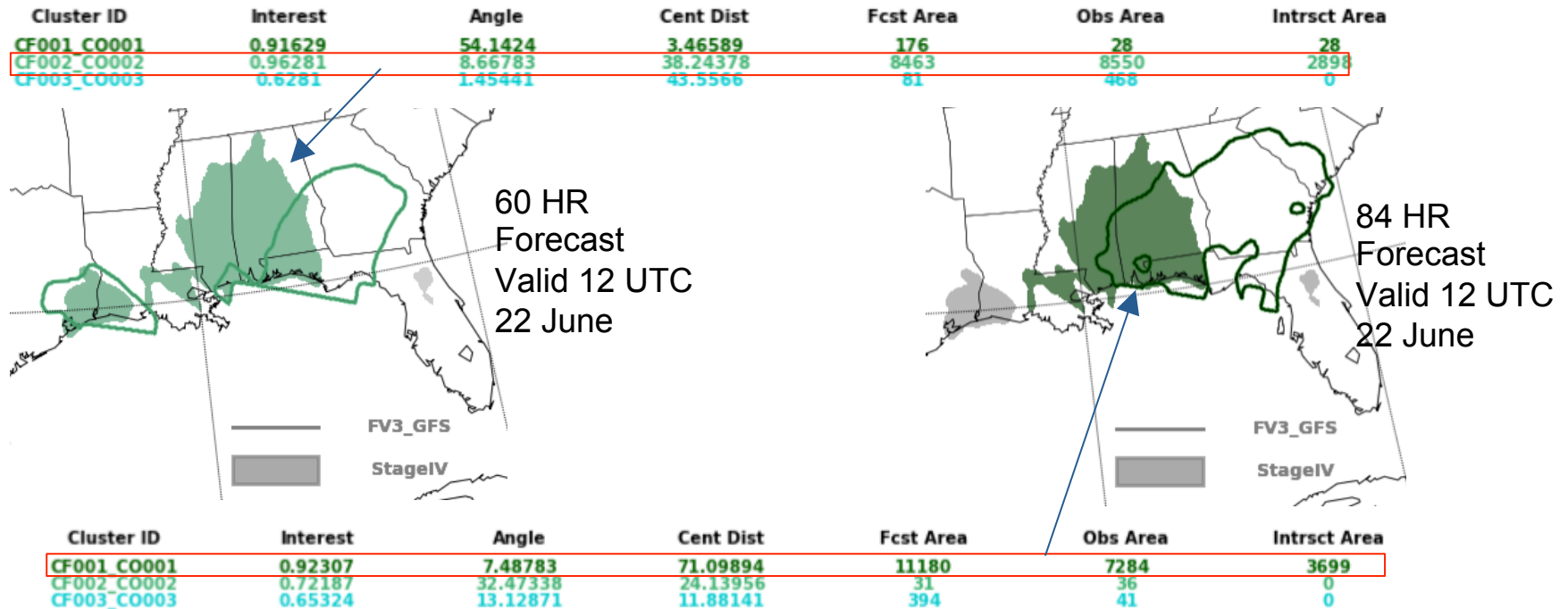


FV3-GFS MODE Examples

FV3_GFS 1" Day 2 24 HR QPF | Initialized: 2017062000 | Valid: 2017061912Z - 20170622 12Z



FV3-GFS MODE Examples



This is an interesting example because traditional grid stats tell us the 84 HR forecast was better (CSI of 0.185 compared to 0.17). However MODE statistics suggest the 60 HR forecast was better for the main precipitation area (lower angle, better centroid distance)

FV3 GFS MODE Analysis

60 HR Forecast
0.5" Threshold
180 Shapes Analyzed

Field	Min	Max	Mean	Std Dev	P25	P50	P75
Centroid Dist	0.64	92.67	22.94	16.41	12.67	18.96	29.72
Angle Diff	0.048	86.14	27.2	23.68	7.57	19.27	42.29
Intersection Area	0	26957	1028.89	2901.86	0	57	581.5
Interest	0.59	1	0.85	0.14	0.69	0.93	0.96

FV3 GFS MODE Analysis

60 HR Forecast
1.0" Threshold
69 Shapes Analyzed

Field	Min	Max	Mean	Std Dev	P25	P50	P75
Centroid Dist	1.65	84.98	24.32	17.09	12.18	20.66	34.27
Angle Diff	0.038	89.25	27.19	25.02	4.73	22.47	42.33
Intersection Area	0	5854	351.97	914.72	0	23	241
Interest	0.54	1	0.83	0.14	0.7	0.85	0.95

FV3 GFS MODE Analysis

84 HR Forecast
0.5" Threshold
155 Shapes Analyzed

Field	Min	Max	Mean	Std Dev	P25	P50	P75
Centroid Dist	0.64	134.33	27.88	21.42	13.86	21.24	35.27
Angle Diff	0.025	89.73	27.2	24.74	8.56	19.02	42.66
Intersection Area	0	21555	1006.34	2844.65	0	48	474.5
Interest	0.57	1	0.83	0.13	0.7	0.9	0.95

FV3 GFS MODE Analysis

84 HR Forecast
1.0" Threshold
62 Shapes Analyzed

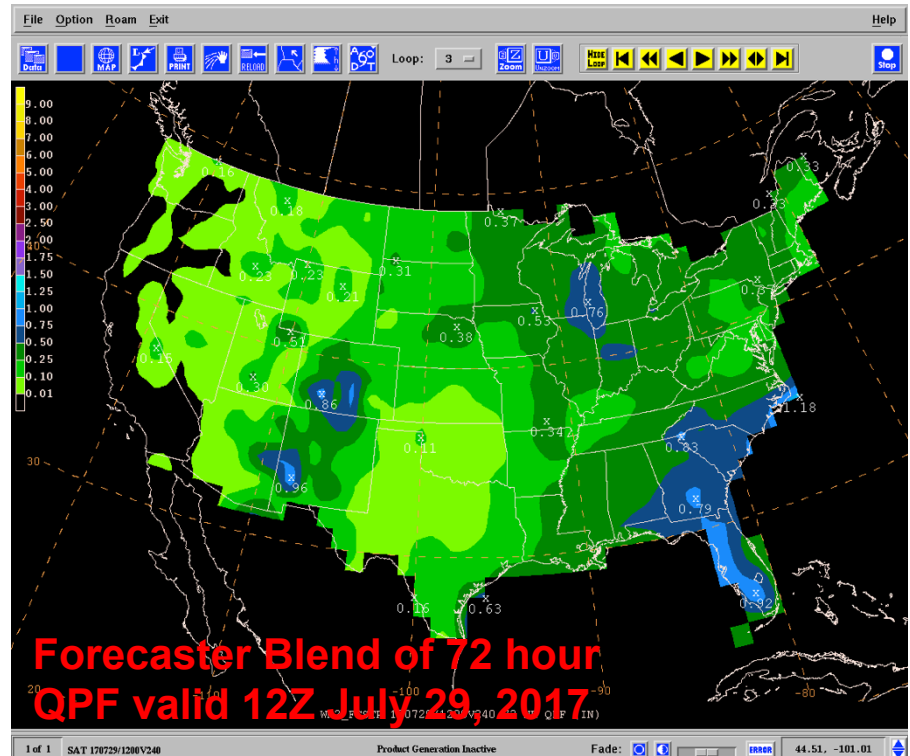
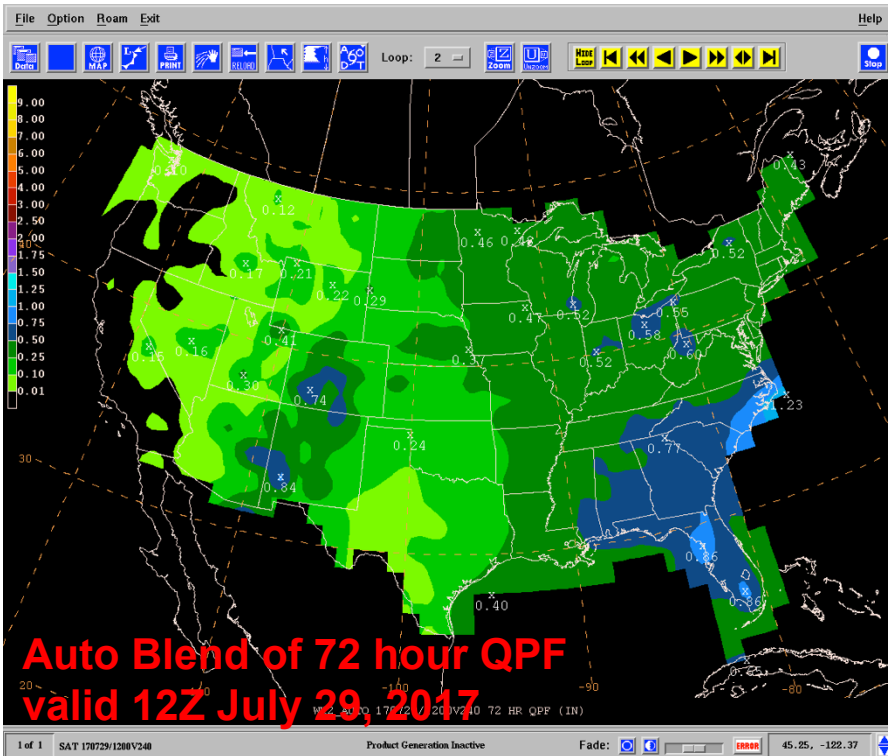
Field	Min	Max	Mean	Std Dev	P25	P50	P75
Centroid Dist	4.76	74.25	30.75	20.04	17.27	24.38	40.3
Angle Diff	0.43	86.96	29.32	25.9	8.59	19.36	49
Intersection Area	0	3793	321.44	774.19	0	31.5	156
Interest	0.61	1	0.81	0.14	0.65	0.83	0.93

DAY 8-10 EXPERIMENT

Use of FV3-GFS in the Day 8-10 Experiment

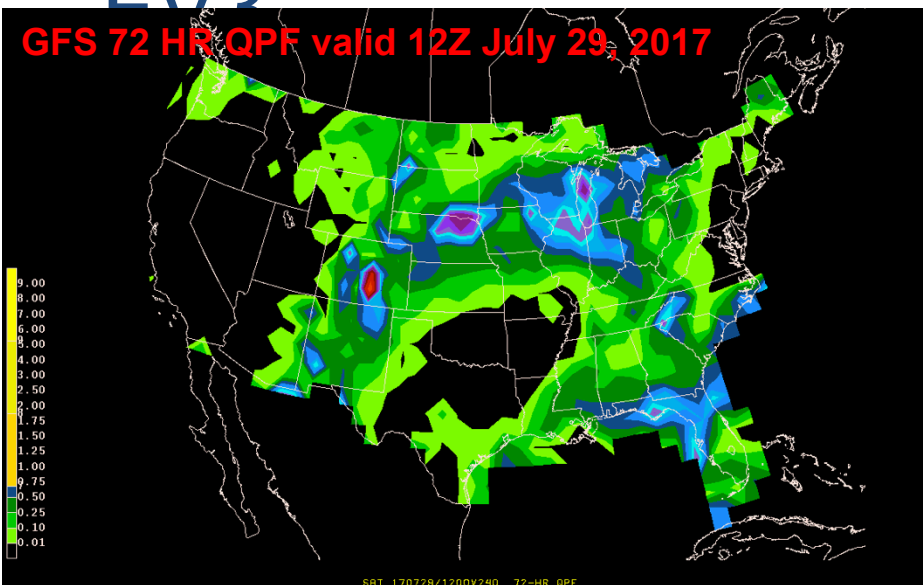
- Auto Blend is the base forecast, consisting of GEFS, ECENS, CMCE, GFS, and GEFS Reforecast.
- Forecasters can then reassign weighting to base models in the blend and add other components i.e. climatology, FV3, and bias corrected GEFS with the goal of improving the forecast.
- QPF available in FV3 and is used in the forecaster blend.
- WPC requests/needs max/min 2-meter temperatures in order to use FV3 in forecast temperature exercises.

Sample QPF for Day 8-10

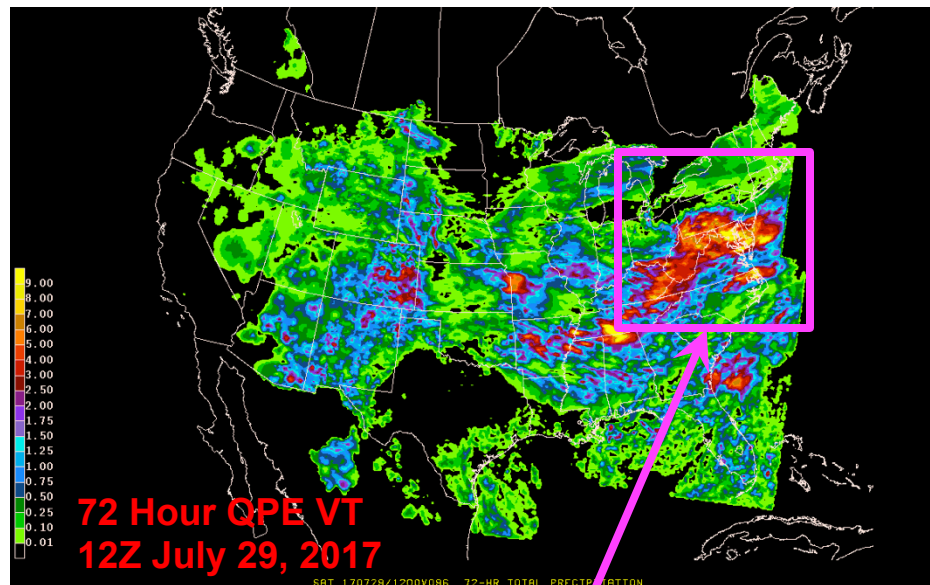


Comparing 72 Hour QPF From GFS and FV3

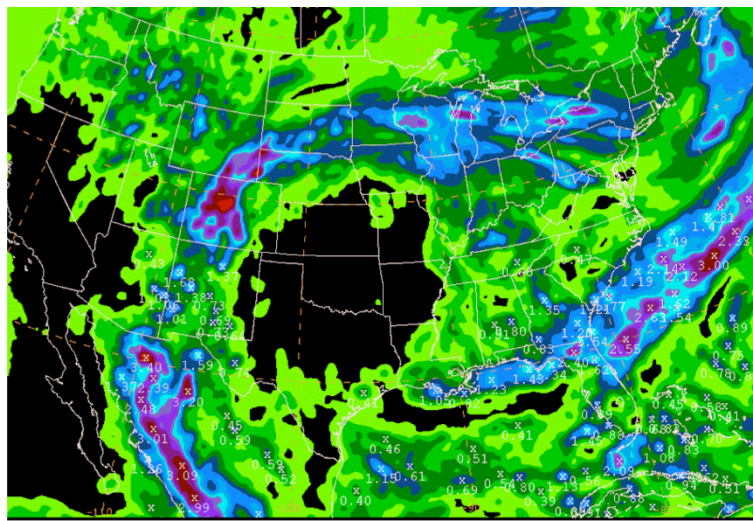
GFS 72 HR QPF valid 12Z July 29, 2017



72 Hour QPE VT
12Z July 29, 2017



FV3-GFS 72 HR QPF VT 12Z July 29, 2017



Difficult to forecast big convection precipitation events in the Day 8-10 period

- Skill much lower with warm season QPF.
- Often significant differences between GFS and FV3
- Will occasionally assign a low weight to the FV3 in addition to the GFS, or substitute for the GFS.

Thank you!

- Questions?
- benjamin.albright@noaa.gov
- joshua.kastman@noaa.gov
- mike.bodner@noaa.gov